Multi-Dimensional Prediction of Guild Health in Online Games: A Stability-Aware Multi-Task Learning Approach

Chuang Zhao[1], Hongke Zhao[1*], Runze Wu[2*], Qilin Deng[2], Yu Ding[2], Jianrong Tao[2], Changjie Fan[2]

[1]College of Management and Economics, Tianjin University,
[2]Fuxi AI Lab, NetEase Games,
{zhaochuang, hongke}@tju.edu.cn, {wurunze1, dengqilin, dingyu01, hztaojianrong, fanchangjie}@corp.netease.com,

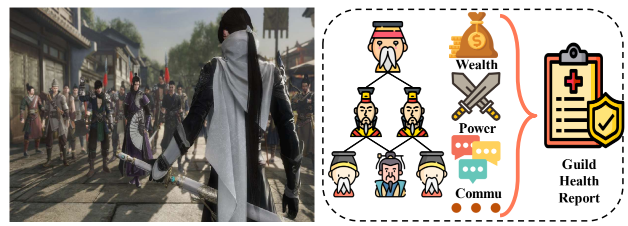
Abstract

Guild is the most important long-term virtual community and emotional bond in massively multiplayer online role-playing games (MMORPGs). It matters a lot to the player retention and game ecology how the guilds are going, e.g., healthy or not. The main challenge now is to characterize and predict the guild health in a quantitative, dynamic, and multi-dimensional manner based on complicated multi-media data streams. To this end, we propose a novel framework, namely Stability-Aware Multi-task Learning Approach (SAMLA) to address these challenges. Specifically, different media-specific modules are designed to extract information from multiple media types of tabular data, time series characteristics, and heterogeneous graphs. To capture the dynamics of guild health, we introduce a representation encoder to provide a time-series view of multi-media data that is used for task prediction. Inspired by well-received theories on organization management, we delicately define five specific and quantitative dimensions of guild health and make parallel predictions based on a multi-task approach. Besides, we devise a novel auxiliary task, i.e., the guild stability, to boost the performance of the guild health prediction task. Extensive experiments on a real-world large-scale MMORPG dataset verify that our proposed method outperforms the state-of-the-art methods in the task of organizational health characterization and prediction. Moreover, our work has been practically deployed in online MMORPG, and case studies clearly illustrate the significant value.

Introduction

Massively multiplayer online role-playing games (MMORPGs) have become a phenomenon of growing cultural, social, and economic importance, routinely attracting millions of players (Liao et al. 2020). Guild, often called tribe or clan in MMORPGs, is a smaller piece of the community made up of users with shared goals (Pisan 2007). It allows players to interact more intimately with the guild's members, which is of great significance to character leveling, player retention, game activity, etc (Poor 2015). However, the guild health will deteriorate due to various reasons, gradually becoming inactive or disbanding directly, which harms both the player and the platform.

*Corresponding Authors



(a) One guild in a famous MMORPGs *Justice Online*. (b) Dimensions of guild health.

Figure 1: Guild health. Guild health can be measured from multiple dimensions.

Considering the guild is a specific organization in MMORPGs, we refer to the existing literature on organization and its health. Traditional research on organizations in the field of management mostly relies on survey interviews and classical statistical models (Meggetto et al. 2020). The interpretability and low-cost characteristics of these methods make them popular with many researchers and have achieved rich academic results. However, these research methods are mainly based on simple static research scenarios, which are inconsistent with the complicated real-time game scenarios we are discussing. Another research paradigm is to characterize the activities and cohesion of the organization from the perspective of social network analysis, and to explore the different network structure attributes of the organization in different life cycle stages (Lee et al. 2013). Note that this method can effectively use the organizational structure information and perform large-scale calculations (Wu and Pan 2018). However, due to the difficulty of obtaining member portraits and the ambiguity of relationship types in traditional company research, most of these methods do not consider node-level attributes and treat the connections between members as homogeneous.

In general, the previous researches mainly suffer from three challenges. Firstly, the organizational structure of guilds is complicated, coupling as well as heterogeneous (Rodríguez et al. 2019). Enriched with different elements for building a virtual world, MMORPGs can track a variety of data of a guild, including the static characteristics of guilds, the portraits of guild members, the heterogeneous relationships among guild members (e.g., teaming, trading, chatting, etc.) and complicated interactions among them. Secondly, the dynamic nature of the guild health has not been well explored. On the one hand, the guild is not a static organization, and the instability of its members will have a great impact on the guild health. On the other hand, the guild health will continue to change over time, and how to characterize and predict the guild health based on the historical health status is an urgent problem to be solved. Thirdly, it is not easy to directly quantify guild health, actually a multi-dimensional measure. Obviously, as shown in Figure 1, guild health usually consists of factors such as wealth, combat power, and communication. It is not advisable to judge the totally guild healthy based on the value of any one dimension. For example, we cannot say that a highly chatty organization is healthier than a militant organization, especially in the game scene.

To address these challenges, this paper proposes a new data-driven end-to-end neural network framework, namely Stability-Aware Multi-task Learning Approach (SAMLA), to solve the problem of calculating the guild health. Firstly, for characterizing the multiple heterogeneous complicated coupling relationship structures, we integrate all the media data related to the guild, which can be divided into three entities: *guild portrait*, *player portrait* and *heterogeneous relationship among players*. In addition to using Wide and Deep (Cheng et al. 2016) and GCN (Gao, Wang, and Ji 2018) to extract different aspects of guild representation, we also design an attention mechanism to capture the mutual influence between guild portraits and guild members. Secondly, since the guild health is dynamic and changeable, the long-term and short-term health of the guild will have a certain impact on future health. We input the guild representation of a period into the Transformer Encoder (Vaswani et al. 2017) to capture the timing information of the guild. Thirdly, for measuring the guild health from multiple dimensions, inspired by the research of management organization (Lyden and Klingele 2000; Zhao et al. 2021), we innovatively divide the guild health into five dimensions, namely *Communication*, *Ability*, *Resource*, *Activity* and *Leadership*. The corresponding specific indicators are illustrated in Appendix *Table 1*. Multi-task learning method is designed to estimate the sub-goals in parallel (Cao et al. 2020; Zhao et al. 2019). As mentioned in (Poor 2015), the movement of guild members is closely related to the prosperity and decline of the guild. For this reason, we take the perception of *guild stability* as an auxiliary task to supplement the prediction of guild health. Besides, since the importance of different dimensions of guild health has not been mentioned in the literature, we design an adaptive weight learner for joint optimization. This approach will automatically generate weights for different tasks from the data and train them together with the model as parameters.

The main contributions of this paper are as follows:

- To our knowledge, this paper represents the first effort to model the health of the guild by integrating multiple heterogeneous complicated coupling media data. A new deep neural network, namely SAMLA, is proposed, which can quantitatively and dynamically characterize and predict the guild health.
- Our method can simultaneously perceive guild stability and predict guild health in a multi-dimensional manner under the framework of multi-task learning. The balance of main and auxiliary tasks is automatically achieved using an adaptive weight learner during joint optimization.
- Intensive experiments have shown that our model performs better than the current state-of-the-art methods for measuring the health of guilds, and can be easily extended to online community research. We also deploy it in several large MMORPG games, which confirms its value in industrial applications.

Related Work

Our proposed method solves the problem of guild health characterization and prediction by constructing a new multimedia neural network framework. Therefore, we briefly return to the most relevant work from two aspects: 1) guild analysis and 2) graph convolutional networks.

Guild Analysis: Guild analysis is a key component of many game researches, such as classification of guild types, exploration of the relationship between guilds and players, and guild life cycle prediction (Liao et al. 2020). Traditional guild analysis is mostly based on simple scale analysis or social network analysis methods. Specifically, Chen, Sun, and Hsieh (2008) divide the guild into five types according to its stability. Pisan (2007) design a scale from the perspective of social identity theory and find that guild identity has a significant impact on player retention. Kang et al. (2015) conduct exploratory factor analysis to analyze the reasons why players join the guild. The result shows that communication and community status play an important role in the guild life cycle. Chung et al. (2014) further define concepts such as group cohesion and group diversity based on social network analysis indicators, and use decision tree methods to obtain the reasons for player churn.

Different from the above studies, we focus on a new research topic, i.e., guild health characterization and prediction, by adopting a new multi-media data-driven neural network approach.

Graph Convolutional Networks: Recently, Graph Convolutional Networks (GCNs) have shown huge success for graph representation learning and related applications (Wu et al. 2020b). Specifically, Sun et al. (2019) improve GCN in their framework to capture the compatibility characteristics perceived by the organization to solve the P-O fit problem. On the identification of high-potential talents in organizations, Ye et al. (2019) combine GCN and social network analysis indicators to construct employee portraits, and achieve good performance on this task. Wu et al. (2020a) design an adaptive GCN framework for both score prediction and user missing attribute inference, and finally achieve better results than a single task.

In this paper, we take advantage of GCN to capture multiple complicated heterogeneous relationships in guilds. To enhance the completeness and accuracy of the guild health characterization and prediction, we also employ a multi-tasking modeling approach, inspired by management work.

Proposed Method

In this section, we first introduce the overall architecture of the proposed model SAMLA and then describe the model optimization process.

Problem Definition

Our multi-media data stream consists of three entities: player portraits, guild portraits, and heterogeneous relationships among players. $V = \{v_1, ..., v_m\}$ is a set of players, where m is the number of players. The player portrait of v_i is denoted as u_i, which is a fixed-length vector. $P = \{p_1, ..., p_n\}$ is a set of guilds, where n is the number of guilds. p_i is made up of $\{v_1, v_2, ..., v_l\}$, where l is the number of players of the guild. The guild portraits of p_i is denoted as x_i, which is a fixed-length vector. Meanwhile, we use $G = (V, E, U)$ as relation graph, where V is the set of players and $U = \{u_1, ..., u_m\}$ denotes the portraits of all players, . In G, E represent all directed edges (w_s, v_i, v_j), where w_s is the weight of the edge and v_i is the player. The weight of the edge is defined as the frequency of the occurrence of the edge within the relation. r is the relationship type, in this paper, $r \in \{trade, friend, team, chat\}$. N_i is player v_i's neighbors. e_v and e_p represent player and guild embedding respectively.

Overview of SAMLA Framework

Figure 2 illustrates the overall architecture of our proposed SAMLA. Specifically, SAMLA includes three main modules: 1) Unified Representation Extraction Module (UREM), 2) Historical State Extraction Module (HSEM), and 3) Guild Health Prediction Module Based on Stability Perception(SHPM). Firstly, media-specific network structures are designed based on multi-media types to embed guild characteristics, player characteristics and diverse relationships among players into a unified representation vector. Multiple representation vectors over a continuous period are then fed into the HSEM to capture the dynamics of the guild health. Finally, inspired by organizational research in management, SHPM innovatively divides organizational health into five specific dimensions while perceiving the stability of guilds, and uses multi-task learning methods to estimate each sub-goal in parallel.

Unified Representation Extraction Module

The media-specific feature extraction module at every moment is mainly composed of three parts, the guild portrait embedding layer, the relationship embedding layer, and the use of attention mechanism to capture the interaction between the guild and the player. To ensure the memorization and generalization of the model, we first use Wide and Deep to encode the original features of the guild p_i at time t. The features of the guild $x_{i,t}$ at t are numerical or categorical.

$$e^W_{p_{i,t}} = WD(x_{i,t}), \tag{1}$$

where $WD(.)$ are Wide and Deep module.

Then use GCN as the main building block to integrate the complicated heterogeneous relationships in the guild structure, and naturally obtain the topological neighborhood structure and edge information of each node. For each player i on each relationship graph r, its representation $h^{l+1}_{i,t}$ on the $l + 1^{th}$ GCN layer is:

$$h^{r,l+1}_{i,t} = \sigma(\sum_j \frac{1}{c^r_{i,j}} h^{r,l}_{j,t} W^l || h^{r,l}_{i,t}), \tag{2}$$

where σ is activation function and W^l is a transformation matrix in the $l^{t,h}$ layer. $c^r_{i,j,t} = \sqrt{d^r_{i,t} d^r_{j,t}}$, where $j \in N_{i,t}$, $d^r_{i,t}$, $d^r_{j,t}$ is the degree of node i, j at time t. $||$ is the concatenation operation. Each player's portrait $u_{i,t}$ is taken as the initial embedding $h^{r,0}_{i,t}$. The final aggregated node representation actually contains information of itself and neighboring nodes, as well as structural features such as edge weight.

Then for each relation graph r, We aggregate players with the same guild ID to obtain guild-level relation representation $e^r_{p_{i,t}}$

$$e^r_{p_{i,t}} = Agg(h^{r,l+1}_{i,t}, \forall i \in N_{p_{i,t}}), \tag{3}$$

where Agg is an aggregate function, here is average pooling. $N_{p_{i,t}}$ are all members of guild p_i at time t and $h^{r,l+1}_{i,t}$ is embedding of the player i in the $l + 1$ layer of GCN.

To further capture the mutual influence between the guild portrait and the guild members, we incorporate $e^W_{p_{i,t}}$ and $e^r_{p_{i,t}}$ into the attention mechanism for updating, as shown in Figure 3. Specificaly, we first process $e^W_{p_{i,t}}$ through the matrix W into the same embedding size as $e^r_{p_{i,t}}$. Then use the self-attention mechanism between the guild portrait and the guild relationship to update the relationship representation.

$$e^W_{p_{i,t}} = W(e^W_{p_{i,t}}), \tag{4}$$

$$e^{r'}_{p_{i,t}} = Concat(MultiHeadAttn(e^r_{p_{i,t}}, e^W_{p_{i,t}})), \tag{5}$$

where $MultiHeadAttn$ is multi head attention and $e^{r'}_{p_{i,t}}$ is a concatenation of all updated relations' embeddings. Then we concat all the relations' embeddings and graph portraits embedding to obtain the unified representation of guild p_i at time t, i.e., $e_{p_{i,t}}$.

$$e_{p_{i,t}} = e^{r'}_{p_{i,t}} || e^W_{p_{i,t}}, \tag{6}$$

Historical State Extraction Module

Due to the excellent performance of Transformer in time series modeling, we use Transformer Encoder to capture the dynamics of the guild health in this section. We first replace the sequence position i with a timestamp t, and then add it to $e_{p_{i,t}}$ to maintain the time order.

$$TE_{(t,2i)} = sin(t/10000^{2i/d}), \tag{7}$$

$$TE_{(t,2i+1)} = cos(t/10000^{2i/d}), \tag{8}$$

$$e^P_{p_{i,t}} = e_{p_{i,t}} + TE_{(t)}, \tag{9}$$

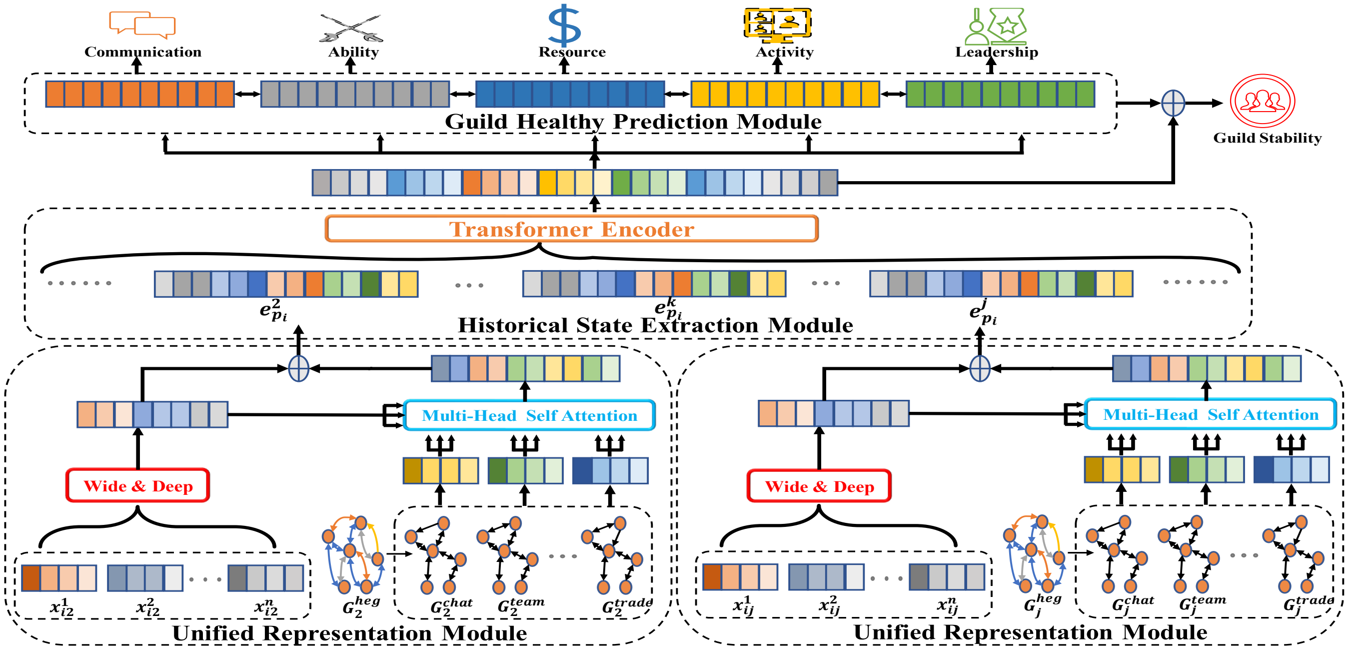


Figure 2: The overall framework of SAMLA.

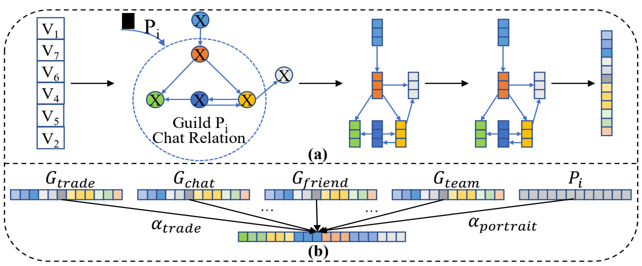


Figure 3: Heterogeneous graph representation extraction. (a) Update node embedding with GCN, and aggregate nodes in p_i with pooling. (b) Attention mechanism, which is used to capture the interaction between the guild and the player.

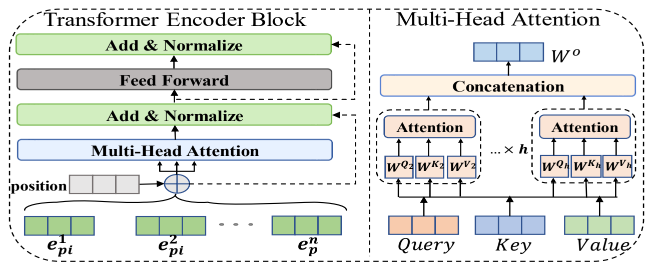


Figure 4: Historical State Extraction Module (HSEM).

where i is the dimension and d is the embedding size. Next, we use the multi-head attention mechanism to capture the long and short-term historical health of the guild.

$$e^M_{p_{i,t}} = MultiHeadAttn(e^P_{p_{i,1}}, ..., e^P_{p_{i,k}}), \tag{10}$$

where k is the length of the history window used for training. Then, we pass the result through several layers to obtain a guild representation with historical health status.

$$e_{p_{i,t}} = LayerNorm(Linear(e^M_{p_{i,t}}) + e^P_{p_{i,t}}), \tag{11}$$

Eventually, all the vectors are concated to form the final unified representation of the guild with historical health $e^H_{p_i}$.

$$e^H_{p_i} = (e_{p_{i,1}}||...||e_{p_{i,k}}), \tag{12}$$

Model Optimization

In this section, we first describe the prediction process of guild health, and then define the loss function of SAMLA for model training.

Guild Health Prediction Module. We design a method to model guild health from multiple dimensions, while using the perception of guild stability as an auxiliary task. Once the guild's unified representation is obtained, it will be split through 5 different linear layer W_i. The five embeddings $e^{dim}_{p_i}$ after splitting are used to predict different dimensions of guild health, namely communication, ability, resources, activity and leadership.

$$e^{dim}_{p_i} = W_i(e^H_{p_i}), \tag{13}$$

where dim is different dimension of guild health, such as communication. Since the health dimensions of guilds are not completely independent, we use parameter sharing mechanism among $e^{dim}_{p_i}$ to capture the interaction among different dimensions. We will illustrate its effectiveness in the experiment. Meanwhile, five segmented embeddings and the unified guild's representations will be concated to perceive the stability of the guild.

$$e^{stability}_{p_i} = e^{dims}_{p_i}||e^H_{p_i}, \tag{14}$$

where $e_{p_i}^{dims}$ is the concatenation of all dimensions' embedding $e_{p_i}^{dim}$ and $e_{p_i}^{stability}$ is used to predict guild stability.
Guild Health Loss. Since the guild health contains 5 dimensions, guild health prediction loss also contains 5 loss values. According to task types, we set loss functions as follows

$$L_{dim} = \frac{1}{n}\sum_{1}^{n}(y_{i,dim} - \hat{y}_{i,dim})^2, \tag{15}$$

where $y_{i,dim}$ is the true value and $\hat{y}_{i,dim}$ is the predicted value of dim. n is the number of all training samples.

Traditional multi-loss optimization method is to perform weighted summation of different loss values, but this method is subjective and not scalable. Therefore, this paper puts the weight as a parameter into the model for learning.

$$L_{health} = \sum_{dim}(\frac{1}{2\sigma_{dim}^2}Loss_{dim} + \log \sigma_{dim}), \tag{16}$$

where σ is noise. σ_{dim} is equivalent to the adaptive weight of loss $Loss_{dim}$ based on data. Because when σ_{dim} increases, the weight of $Loss_{dim}$ decreases, vice versa. The last item $\log \sigma_{dim}$ acts as a noise regulator, suppressing excessive increases in noise. Note that this adaptive weight learner is also suitable for other joint optimization tasks.
Stability Perception Loss. The perception of guild stability is an auxiliary task. Intuitively, unhealthy guilds are usually unstable. Using guild stability as an auxiliary task can supplement information for the guild health prediction task. To further capture the dynamics of the guild stability, we not only consider the original size of the guild but also take into account the number of people who join and leave the guild.

$$y_{stability} = F(join, leave, size), \tag{17}$$

where F is a combination of three variables. Here, F is the sum function. In this task, we also use the mean square error as the loss value.
Total Loss. To perceive the stability of the guild when predicting the guild health, we combine the two loss values, the total loss function is as follows

$$L_{total} = Loss_{health} + \gamma Loss_{stability}, \tag{18}$$

where γ is also an automatically learned parameter to control the perception of guild stability. The detailed algorithm for training is illustrated in Appendix *Algorithm*.

Experiments

In this section, we conduct experiments on a real-world large scale game dataset to evaluate the performance of SAMLA on guild health characterization and prediction.

Experimental Settings

Datasets. We use a dataset from a popular MMORPG game, released by a well-known game company.In total, we have recorded more than 10 billion game logs which include twenty thousand guilds and their portraits, personal information of more than 20 million players, and 4 types of relationship information. Please refer to Appendix *Details of Dataset* for the detailed fields contained in the data set. Code is avaliable at https://github.com/Data-Designer/AAAI2022-GHP.git.
Baselines Methods. We compare our SAMLA model with the following state-of-art baselines for guild health characterization and prediction. We have fine-tuned all the following benchmarks to achieve the best performance.

- **Classical machine learning models:** We use GBRT (Lucchese et al. 2017) and Random Forest(RF) (Breiman 2001). Since these methods are not designed for processing sequential data, we concat sequential vector into one feature as training data.
- **CNN (Liu et al. 2017):** CNN is also widely used to process sequence data. In our experiment, Conv1D is used to extract guild portraits.
- **LSTM (Tay et al. 2017):** This is a classical RNNs network, which using gating mechanism to capture the long-term dependence of the sequence.
- **TCN (Bai, Kolter, and Koltun 2018):** TCN uses the structure of 1-D FCN and causal convolution to make it more capable of retaining historical information.
- **Transformer (Vaswani et al. 2017):** Transformers uses the attention mechanism to replace the RNN structure.
- **GCN (Gao, Wang, and Ji 2018):** It is designed for extracting information from graphs.
- **SAMLA-Single:** This is a variant of SAMLA, it only predicts one task at a time instead of multi-task learning.

Evaluation Metrics. To evaluate the performance of the guild health prediction, we use the popular measurement standards, i.e., mean absolute error (MAE) and root mean square error (RMSE). Smaller values of MAE and RMSE indicate better performance. For other training parameter information, please refer to Appendix *Parameter Settings*.

Comparisons with Baseline Methods

Table 1 shows the performance of our model comparing to the baselines. For fairness, we have all baselines with the same embedding size in accordance with SAMLA model. Obviously, the proposed SAMLA is superior to all baseline methods in any metrics, proving the effectiveness of our framework. Specifically, SAMLA improves over the strongest baseline with 7%, 8%, 18%, 1%, and 3% on Communication, Ability, Resource, Activity, and Leadership. From the results, the performance is much worse for classical methods, such as GBRT and RF. One possible reason is that they are not suitable for handling complicated multimedia data and ignore the dynamics of guild health, resulting in a large amount of information loss. CNN and LSTM surpasses traditional machine learning, which benefits from more complicated modeling. TCN and Transformer donnot show their superiority. This may be due to the short time window and cannot show their superiority in capturing long sequences. Traditional time series prediction algorithms are all still weaker than GCN, which shows that the multiple, complicated, and heterogeneous relationships among guild members are an indispensable part of measuring the guild health. Although GCN has achieved good results, some information is still missing due to its inability to model the

Model	Communication		Ability		Resource		Activity		Leadership		Stability	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
GBRT	74.442	39.979	121.758	38.430	53.554	23.833	69.140	41.291	26.002	15.903	16.470	6.352
RF	70.980	38.739	115.147	41.531	50.610	24.157	67.807	38.809	25.595	15.989	16.072	5.592
CNN	35.426	5.090	66.340	6.990	55.937	7.010	34.337	4.940	31.385	3.840	18.655	2.350
LSTM	33.226	4.680	55.163	4.650	32.388	5.130	32.588	4.680	7.141	0.890	7.348	0.980
TCN	36.878	5.500	70.626	8.10	59.783	7.260	34.871	5.190	29.749	3.540	15.010	2.030
Transformer	35.270	4.300	62.657	6.050	59.034	9.750	34.424	4.340	40.595	7.040	24.454	4.100
GCN	10.607	0.466	11.883	0.536	7.225	0.322	11.507	0.499	6.620	7.869	16.796	0.779
SALMA-Single	8.171	0.302	9.134	0.368	6.729	0.305	11.000	0.441	0.945	0.040	4.595	0.171
SALMA	**7.590***	**0.248***	**8.412***	**0.360***	**5.502***	**0.286***	**10.909***	**0.438***	**0.919***	**0.038***	**2.230***	**0.097***

* means statistically significant by using paired-t test ($p < 0.05$).

Table 1: Performance comparison for guild health characterization and prediction.

long-term health of guilds. We also guess that no perception of guild stability has a certain effect on GCN's results. At the same time, the results of SAMLA are better than those predicted by SAMLA-Single, which fully proves the superiority of multi-task learning method.

Effects of Multi-media Data Streams

In this experiment, different ways of integrating multi-media data streams are replaced to evaluate the effect of its variants.

- **SAMLA-NP:** We hide the player's information in GCN, i.e., ignore the node-level attribute.
- **SAMLA-RA:** In this variant, the attention mechanism is only used among different relationships, i.e., the guild portrait is not used to update the relationship embedding.
- **SAMLA-NA:** When predicting the values of different health dimensions, the parameters of different dimensions $e_{p_i}^{dim}$ are not shared, that is, the interaction among different health dimensions is neglected.

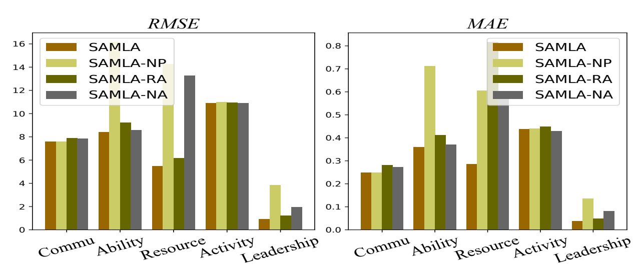


Figure 5: The impact of multi-media data flow influence.

It can be seen from Figure 5 that each media-specific module represents information about different aspects of the guild, and none of them are indispensable. In general, the SAMLA-NP algorithm has the worst performance due to the loss of players' information when obtaining the unified representation of the guild. SAMLA-RA uses different information integration mechanisms, but their performance is weaker than SAMLA, which proves that the mutual effects between the guild portrait and the players cannot be ignored. SAMLA-NA performs slightly worse than SAMLA, which is mainly due to the lack of supplementary information between different health dimensions. Intuitively, the parameter sharing mechanism not only reduce the number of parameters, but also allow information on different health dimensions to complement each other to improve the prediction performance of a single task.

Effects of Perceived Guild Stability

In this section, we substitute other indicators for guild stability to evaluate how the design of guild stability affects the performance of guild health characterization and prediction.

- **SAMLA-NS**: This variant directly predicts the five dimensions of the guild's health, without using the guild stability as an auxiliary task.
- **SAMLA-ND**: This variant replaces the guild stability perception with the size of guild, that is, regardless of the dynamics of the movement of guild members.
- **SAMLA-NT**: This variant directly treats the guild stability prediction as a subtask for training.

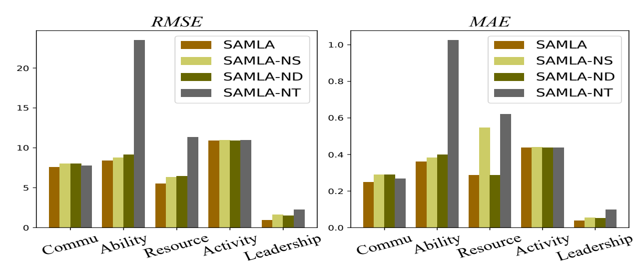


Figure 6: The impact of perceived guild stability.

According to Figure 6, we can get the following observations. First, by comparing SAMLA-NS and SAMLA-ND, it is easy to conclude that the perception of guild stability has a significant impact on the characterization and prediction of guild health. Second, since using the guild size instead of the guild stability ignores that the guild is a dynamically changing structure, the result of SAMLA-ND is slightly worse than our proposed model. Third, SAMLA-NT is the worst variant in most of the dimensional predictions, and we guess the possible reason is that guild stability is not only composed of the five dimensions of guild health. Directly using the guild stability as a subtask will bring the noise to the original main task, namely guild health prediction. In contrast, SAMLA can reduce the impact of this noise with the help of unified representation.

Impact of Hyper-parameters

To verify the impact of different hyper-parameters on SAMLA, we conduct a set of experiments with varied hyper-parameters on our dataset.

Encoder Size. The encoder size determines the high-level characterization extraction capability. We test the following variants of SAMLA model with the encoder size of $1, 2, 3, 4$: **SAMLA-E1, SAMLA-E2, SAMLA-E3** and **SAMLA-E4**.

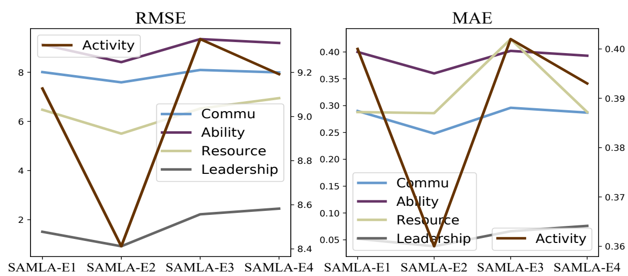


Figure 7: The impact of Encoder size.

As shown in Figure 7, the SAMLA achieves the highest performance when the encoder size is set to 2. Fewer encoders will lead to insufficient model characterization capabilities, and more encoders will lead to over-parameterization of the model (Vaswani et al. 2017).

The GCN Layer Depth. The quality of player representation is closely related to the guild representation, and the number of GCN layers controls the depth of the model and the degree of player aggregation. We test our model with different numbers of layers of $1, 2, 3, 4$: **SAMLA-G1, SAMLA-G2, SAMLA-G3** and **SAMLA-G4**.

As shown in Figure 8, stacking two GCN layers can obtain the best results. Fewer layers cannot capture high-hop neighbor information, and more layers will cause over-smoothing problems (Wu et al. 2020b).

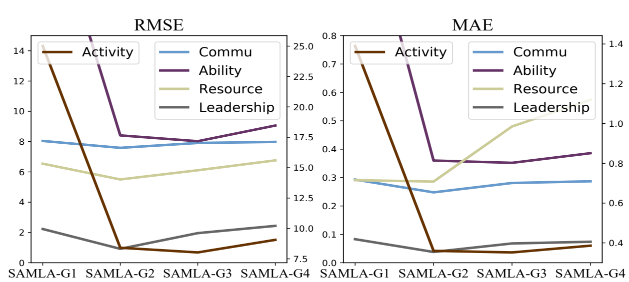
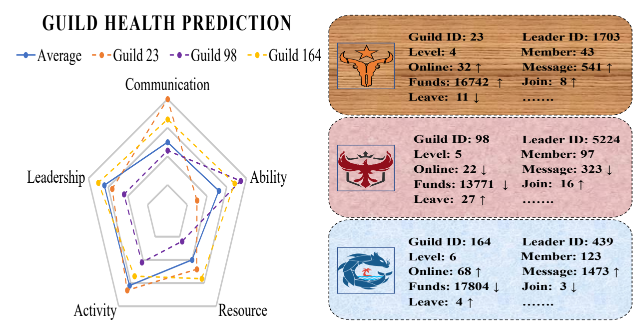


Figure 8: The impact of GCN Layer Depth.

Case Study

In the deployment, we use two forms of graphs to help operators propose targeted measures against the guilds based on the predicted results. We randomly select several typical cases from the service for detailed analysis, and the corresponding prediction results are shown in Figure 9. For example, although the ability dimension of Guild 98 is expected to be better, it is expected to be lower than average in terms of resources and activities. Operators can take some incentive measures in advance, such as increasing the value of guild activities rewards or sending SMS recalls. Another example is guild 23. Its lack of ability dimension is the main reason for its unhealthy. The operator can suggest that the guild absorb more senior talents, or organize multiple battles to improve the overall ability. These targeted measures are of great significance for maintaining the prosperity of the guild and increasing the mobility and retention of players.



(a) Comparison of the average health of all guilds in the server.

(b) Several typical cases.

Figure 9: Real-world application.

Conclusion

In this work, we explore a new task of characterizing and predicting the guild health in multiplayer online role-playing games. This study fills the gap of exploring data-driven learning method for organizational health in a quantitative, dynamic, and multi-dimensional manner. Specifically, we first design different media-specific networks to extract information from various aspects of the guild in a dynamic manner. Then, inspired by many well-received theories on organization management, we delicately define five specific and quantitative dimensions of guild health and apply the multi-task learning approach for parallel prediction. Finally, we innovatively use guild stability as an auxiliary task, which is proved to be an informative supplement to the guild health prediction task. We also conduct a case study to prove the great value of SAMLA in business. In general, the guild health prediction brings convenience to the management of the guilds, improves the game experience of players, and maintains a good game ecology. For future work, we plan to extend our model in other areas, such as online communities and real-world organizational research.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China (72101176, 71790590/71790594, 72091210/72091214).

References

Bai, S.; Kolter, J. Z.; and Koltun, V. 2018. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.

Breiman, L. 2001. Random forests. *Machine learning*, 45(1): 5–32.

Cao, J.; Wang, Y.; He, J.; Liang, W.; Tao, H.; and Zhu, G. 2020. Predicting Grain Losses and Waste Rate Along the Entire Chain: A Multitask Multigated Recurrent Unit Autoencoder Based Method. *IEEE Transactions on Industrial Informatics*, 17(6): 4390–4400.

Chen, C.-H.; Sun, C.-T.; and Hsieh, J. 2008. Player guild dynamics and evolution in massively multiplayer online games. *CyberPsychology & Behavior*, 11(3): 293–301.

Cheng, H.-T.; Koc, L.; Harmsen, J.; Shaked, T.; Chandra, T.; Aradhye, H.; Anderson, G.; Corrado, G.; Chai, W.; Ispir, M.; et al. 2016. Wide & deep learning for recommender systems. In *Proceedings of the 1st workshop on deep learning for recommender systems*, 7–10.

Chung, T.; Han, J.; Choi, D.; Kwon, T. T.; Kim, H. K.; and Choi, Y. 2014. Unveiling group characteristics in online social games: a socio-economic analysis. In *Proceedings of the 23rd international conference on World wide web*, 889–900.

Gao, H.; Wang, Z.; and Ji, S. 2018. Large-scale learnable graph convolutional networks. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1416–1424.

Kang, A. R.; Park, J.; Lee, J.; and Kim, H. K. 2015. Rise and fall of online game groups: Common findings on two different games. In *Proceedings of the 24th International Conference on World Wide Web*, 1079–1084.

Lee, H. J.; Choi, J.; Kim, J. W.; Park, S. J.; and Gloor, P. 2013. Communication, opponents, and clan performance in online games: a social network approach. *Cyberpsychology, Behavior, and Social Networking*, 16(12): 878–883.

Liao, G.-Y.; Pham, T. T. L.; Cheng, T.; and Teng, C.-I. 2020. How online gamers' participation fosters their team commitment: Perspective of social identity theory. *International Journal of Information Management*, 52: 102095.

Liu, J.; Chang, W.-C.; Wu, Y.; and Yang, Y. 2017. Deep learning for extreme multi-label text classification. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 115–124.

Lucchese, C.; Muntean, C. I.; Nardini, F. M.; Perego, R.; and Trani, S. 2017. Rankeval: An evaluation and analysis framework for learning-to-rank solutions. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1281–1284.

Lyden, J. A.; and Klingele, W. E. 2000. Supervising organizational health. *Supervision*, 61(12): 3–6.

Meggetto, E.; Kent, F.; Ward, B.; and Keleher, H. 2020. Factors influencing implementation of organizational health literacy: a realist review. *Journal of health organization and management*.

Pisan, Y. 2007. My guild, my people: Role of guilds in massively multiplayer online games. In *Proceedings of the 4th Australasian conference on Interactive entertainment*, 1–5.

Poor, N. 2015. What mmo communities don't do: A longitudinal study of guilds and character leveling, or not. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 9.

Rodríguez, I.; Kozusznik, M. W.; Peiró, J. M.; and Tordera, N. 2019. Individual, co-active and collective coping and organizational stress: A longitudinal study. *European Management Journal*, 37(1): 86–98.

Sun, Y.; Zhuang, F.; Zhu, H.; Song, X.; He, Q.; and Xiong, H. 2019. The impact of person-organization fit on talent management: A structure-aware convolutional neural network approach. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1625–1633.

Tay, Y.; Phan, M. C.; Tuan, L. A.; and Hui, S. C. 2017. Learning to rank question answer pairs with holographic dual lstm architecture. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval*, 695–704.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention is all you need. *arXiv preprint arXiv:1706.03762*.

Wu, L.; Yang, Y.; Zhang, K.; Hong, R.; Fu, Y.; and Wang, M. 2020a. Joint item recommendation and attribute inference: An adaptive graph convolutional network approach. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 679–688.

Wu, P.; and Pan, L. 2018. Mining application-aware community organization with expanded feature subspaces from concerned attributes in social networks. *Knowledge-Based Systems*, 139: 1–12.

Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; and Philip, S. Y. 2020b. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*.

Ye, Y.; Zhu, H.; Xu, T.; Zhuang, F.; Yu, R.; and Xiong, H. 2019. Identifying high potential talent: A neural network based dynamic social profiling approach. In *2019 IEEE International Conference on Data Mining (ICDM)*, 718–727. IEEE.

Zhao, H.; Jin, B.; Liu, Q.; Ge, Y.; Chen, E.; Zhang, X.; and Xu, T. 2019. Voice of charity: Prospecting the donation recurrence & donor retention in crowdfunding. *IEEE Transactions on Knowledge and Data Engineering*, 32(8): 1652–1665.

Zhao, H.; Liu, X.; Zhang, X.; Wei, Y.; and Liu, C. 2021. The effects of person-organization fit on lending behaviors: Empirical evidence from Kiva. *Journal of Management Science and Engineering*.